HAL open science

Improvement of ENSO prediction using a linear regression model with a Southern Indian Ocean Sea Surface Temperature Predictor

Sébastien Dominiak, Pascal Terray

▶ To cite this version:

Sébastien Dominiak, Pascal Terray. Improvement of ENSO prediction using a linear regression model with a Southern Indian Ocean Sea Surface Temperature Predictor. Geophysical Research Letters, 2005, 32, pp.L18702. 10.1029/2005GL023153 . hal-00124054

HAL Id: hal-00124054
https://hal.science/hal-00124054

Submitted on 11 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Improvement of ENSO prediction using a linear regression model with a southern Indian Ocean sea surface temperature predictor

Sébastien Dominiak[1] and Pascal Terray [1,2]

[1] *Laboratoire d'Océanographie Dynamique et de Climatologie, IPSL, Paris, France*

[2] *Université Paris 7, Paris, France*

Submitted to ***Geophysical Research Letters,*** *Revised* June 2005

authors address :

Sébastien Dominiak (email: **sdolod@lodyc.jussieu.fr**)

Pascal Terray (email: **terray@lodyc.jussieu.fr**)

Laboratoire d'Océanographie Dynamique et de Climatologie, IPSL, Université Pierre et Marie Curie,

BP 100, 4 Place Jussieu, 75252 Paris CEDEX 05, France

Abstract

[1] This study presents a detailed comparison between three ENSO precursors which can predict across the spring persistence barrier: the anomalous equatorial Pacific upper ocean heat content, the zonal equatorial wind stress anomaly in the far-western Pacific and SST anomalies in the South-East Indian Ocean (SEIO) during the late boreal winter. A new correlation analysis confirms that El Niño (La Niña) onsets are preceded by significant cold (warm) SST anomalies in the SEIO during the late boreal winter after the 1976-77 climate regime shift. Thus, the objective is to examine the respective potential of these three ENSO precursors to predict ENSO events across the boreal spring barrier during recent decades. Surprisingly, in this focus, cross-validated hindcasts of the linear regression models based on the lagged relationship between Niño3.4 SST and the predictors suggest that SEIO SST anomalies during the late boreal winter is the more robust ENSO predictor.

1. Introduction

[2] Predicting the occurrence the El Niño-Southern Oscillation (ENSO) phenomenon several months in advance is a major goal of climate research. During recent years, dynamical models have been used to this end, but despite great efforts and significant progress, they encounter many difficulties to forecast the timing, structure and amplitude of ENSO events [Landsea and Knaff, 2000]. On the other hand, statistical models show modest, but still significant skill in forecasting various ENSO indices such as Niño3.4 Sea Surface Temperature (SST) [Clarke and Van Gorder, 2003]. In this study, we discuss a new ENSO precursor, SEIO SSTs during late boreal winter, and demonstrate through various linear regression exercices that this new parameter leads to a significant improvement in ENSO statistical prediction across the so-called boreal spring persistence barrier during recent decades [Torrence and Webster, 2000].

2. A selection of robust ENSO precursors.

[3] Two critical factors have been identified to have a close relationship with the onset of El Niño events and are pivotal in many ENSO prediction exercises [Clarke and Van Gorder, 2003]. First, the equatorial Pacific upper ocean heat content anomaly during the boreal winter preceding El Niño onset plays an important role in ENSO evolution [Meinen and McPhaden, 2000]. Second, the local wind anomalies over the western Pacific and Indian Ocean warm pool region have also been considered as a cause for ENSO variability [McPhaden et al., 1998]. Both are two excellent ENSO predictors that can predict across the boreal spring [Clarke and Van Gorder, 2003]. Drawing on the experience of two very recent studies [Terray et al., 2005, Terray and Dominiak, 2005], a further predictor is introduced here, SEIO SSTs during boreal winter. The two aforementioned studies suggest that SEIO SST anomalies have

a delayed and/or prolonged impact on ENSO variability through two mechanisms: an external forcing of local wind anomalies over the Pacific warm pool and a modulation of the regional Hadley cell in the southwest Pacific [Terray and Dominiak, 2005]. Several other very recent studies have also confirmed that SST and convective anomalies over the eastern Indian Ocean are physically linked to the wind anomalies over the western Pacific [Kug et al., 2005]. These findings suggest that the Indian Ocean could be one more important factor in contributing to ENSO evolution during the recent decades.

[4] Figure 1 shows the lag correlations between December-January Niño3.4 SSTs and the previous February-March surface zonal wind [from NCEP/NCAR reanalysis, Kalnay et al., 1996], 20°C isotherm depth [from the SODA data, Carton et al., 2000] and SST [from the Extended Reconstruction of global SST, Smith and Reynolds, 2004] over the Indo-Pacific area. Computations are made for the 1977-2001 period, to avoid possible effects due to the so-called 1976-77 climate regime shift [Nitta and Yamada, 1989], as it is known that ENSO precursors changed significantly around this date [Wang, 1995]. This correlation analysis illustrates the significance of the two ENSO precusors introduced by Clarke and Van Gorder [2003], the western equatorial Pacific zonal wind (130°E-160°E, 5°S-5°N, black frame in Figure 1a, UWPAC hereafter) and the equatorial Pacific upper ocean heat content (130°E-80°W, 5°S-5°N, black frame in Figure 1b, hereafter Z20). Consistent with the ENSO recharge-oscillator paradigm [Jin, 1997], the mean depth of the February-March 20°C isotherm over the equatorial Pacific is positively and significantly correlated (0.72) with Niño3.4 SSTs ten months after. This suggests that the amount of warm water in the tropical Pacific builds up prior to El Niño onset. Thus, a substantial heat content anomaly in the Pacific Ocean mixed layer seems to be a necessary contition for El Niño growth. The significant positive correlation (0.75) between February-March surface zonal wind over the

western Pacific Ocean warm pool region and December-January Niño3.4 SSTs also confirms that the wind anomalies in this area is an important factor for ENSO variability (Figure 1a). These westerly wind anomalies can induce eastward propagating equatorial Kelvin waves and have also been proposed to be a trigger for some El Niño events. Finally, the highest correlations with the SST fields are found in the Indian Ocean (-0.77). The SEIO area (90°E-122°E, 5°S-45°S, black frame in Figure 1c) seems to be another strategic factor for determining of developing El Niño or not before the onset [Terray and Dominiak, 2005]. Thus, this precursor is also a good candidate for ENSO prediction across the boreal spring barrier.

3. Linear prediction results

[5] We assess now the usefulness of SEIO SSTs to improve the limited amount of skill of current statistical models in predicting ENSO from the late boreal winter season (February-March) before the El Niño onset. To this end, we developed various regression models for forecasting Niño3.4 SST anomalies for different time averages (October-December, November-January, December-February) around the ENSO peak season, using only February-March anomalies of the predictors. As a measure of the predictive skill, we compare the observed Niño3.4 SST with the values calculated from regression equations based successively on all years with the 1977-2001 time span, except the forecast year. To assess the forecast potential, three statistics have been used: the correlation coefficient (COR) between the observed and forecast Niño3.4 SST, the Root-Mean-Square-Error (RMSE) and the so-called Performance Parameter (PP):

$$PP = 1 - (RMSE/SD)^2 \quad (1)$$

where SD is the standard-deviation of Niño3.4 SST for the respective season. When PP is greater (lower) than zero the forecast is better (worse) than a climatological forecast.

[6] We start by discussing the three simple linear regression models where each predictor serves as sole imput (Table 1). For the October-December season, the SEIO and UWPAC models yield virtually the same RMSE and correlation coefficient between predicted and observed Niño3.4 SSTs, with correlation coefficients of 0.7 and RMSE of 0.68, while the forecast performance of Z20 is inferior using this cross-validated method. For longer time leads, SEIO exhibits consistently higher correlations and lower RMSEs than Z20 and UWPAC. Surprisingly, the predictive skills for SEIO precursor are also better when the time lead between the predictor and predictand is increased. Finally it is interesting to note that all of these univariate models perform better than a climatological forecast since PP is always greater than zero.

[7] Figure 2 shows the observed and forecast October-December Niño3.4 SST for all the univariate models. Each predictor exhibits successful forecasts, but also major failures. The Z20 model has a high predictive skill for La Niña events, but this model has major failures for predicting El Niño events, excepted the 1997-98 El Niño. Interestingly, this latter event is very well-captured by the UWPAC model. This is consistent with results stressing the crucial importance of the March 1997 westerly wind burst over the Pacific warm pool for the onset of the 1997-98 El Niño [Wang and Weisberg, 2000]. However, the performance of this last model for other ENSO events is rather poor. The SEIO model has a high predictive skill in the first half of the record, while this model has a substantially degraded performance for the more recent period. Finally, all three univariate models have major difficulties in the 1990s

when ENSO events are known to be particularly unusual. During the 1991-94 time interval, a substantial warming is observed in the central Pacific and even if these years correspond to one or several events is a controversial matter [Trenberth and Hoar, 1996; Van Loon et al., 2003]. These results serve to emphasize that a complicated combination of atmospheric and oceanic conditions are necessary for El Niño to occur [McPhaden et al., 1998].

[8] A natural way of improving the univariate models is to add new variables (table 2). Comparison of Table 2 with Table 1 bears out the overall remarkably superior performance of the updated models. Note that SEIO remains the most important predictor since the models have a higher predictive skill over nearly all the forecast lead times when SEIO is included in the model. Nevertheless, the other two elements add useful information which always improve the performance skill of the univariate SEIO model (cf. Table 1). It appears clearly that the most efficient model is built using SEIO and UWPAC. Regardless of which lead time is chosen, the same hierarchy is found with this model performing somewhat better than the others, followed by Z20-SEIO model and afterwards Z20-UWPAC. This result confirms that SEIO SSTs are not only a precursor of Niño 3.4 SSTs via the zonal wind anomalies over the far western equatorial Pacific, but also via a modulation of the regional Hadley cell in the Southwest Pacific ocean [Terray and Dominiak, 2005]. Moreover, this regression exercise also suggests that the Z20 precursor is of subordinate importance in predicting ENSO variability from the late boreal winter setting.

[9] Figure 3 shows cross-validated Niño3.4 SST time series forecast for November-December for the “two predictors” models. Considering the UWPAC-SEIO model, we observe that many El Niño or La Niña events are correctly identified, especially in the 1980s when this model predicts the occurrence and amplitude of ENSO events. The very strong 1997-98 El

Niño is also identified, but with a weaker amplitude. Discrepancies between the forecast and observed values for the UWPAC-SEIO model correspond to some consecutive years of anomalously warm or cold Niño3.4 SSTs, such as 1986-87 or 1991-94. For this last time interval, Goddard and Graham [1997] suggest that the 1993-94 event is partly due to the persistence of warm SSTs in the central equatorial Pacific, these features being not linked to the selected precursors. Now considering the Z20–SEIO model, we observe a high predictive skill of La Niña events. This is consistent with the behaviour of these two predictors in an univariate setting, as seen on Figure 2. On the other hand, this model performs poorly on El Niño events. Consistent with the work of Wang and Weisberg [2000], the Z20-UWPAC model predicts the very strong 1997-98 El Niño with the highest predictive skill in all the models presented in this study. However, the performance of this model is largely degraded for the rest of the record. In synthesis, it appears that UWPAC-SEIO model is surely the best one in predicting ENSO variability. But other models are able to lead to better predictions on specific occasions, again illustrating the complexity of the Pacific climate system.

[10] Surprisingly, the Z20-UWPAC-SEIO model does not lead to any significant improvement in predictive skill whatever the time lead may be (Table 3). This model exhibits a similar efficiency as the UWPAC-SEIO model. This raises again the question of the importance of the Z20 precursor during recent decades. Does one need to consider that Z20 is a crucial parameter for ENSO prediction because of its physical relevance [Jin, 1987; Meinen and McPhaden, 2000] ? This question needs further exploration. First, the weakness of the SODA data is a possible factor for explaining the present results. In order to check the reliability of the SODA reanalysis, the Z20 time series has been recomputed from the BMRC ocean thermal analysis [Smith, 1995] which covers the period 1980-2001. The correlation coefficient between the SODA and BMRC February-March Z20 time series is as high as 0.93

during the 1980-2001 time span and is significant at the 99% confidence level. The possibility that the results may improve with the BMRC analysis has been further examined by performing additional cross-validated regression experiments where the Z20 time series is computed from the BMRC data. The statistics for these experiments are reported in Table 4. It is noteworthy that the predictive skills of the experiments using the BMRC index are not significantly better than the skills obtained from the models using the SODA reanalysis with similar correlation coefficients, but higher RMSEs. This suggests that the poor performance of the Z20 parameter is not due to a deficiency of the SODA data. Thus, according to the principle of parsimony which requires that a statistical model should employ a number of predictors as small as possible, the multiple linear regression model with all three precursors as input offers no advantage. Of course, this conclusion is only valid for the specific time leads examined here. For longer time leads (eg for more than one year), Z20 remains a highly significant precursor (not shown) since the transition for SEIO SST anomalies occurs in early boreal winter when the seasonal wind over the eastern Indian Ocean turns from southeasterly to northwesterly (Terray et al., 2005).

4. Concluding remarks.

[11] This study explores the robustness of ENSO precursors which have not been previously considered. In addition to well-recognized precursors of El Niño onsets, such the Z20 and UWPAC parameters during the late boreal winter, a further predictor is examined here, namely SST anomalies in the SEIO. Various statistical prediction models for predicting Niño3.4 SST were constructed based on the lag relationship between Niño3.4 SST and these three predictors observed during the late boreal winter. Surprisingly, the most prominent precursor is SEIO SSTs, whereas the much less powerful precursor is the Z20 parameter. The

forecast skill is considerably better for the models which use SEIO SST as a predictor over different lead times during the 1977-2001 period.

[12] Considering the best model, the UWPAC-SEIO model, the highly significant correlations and PP scores between the forecast and observed Niño3.4 SST index indicate that about half of the interannual variability of Niño3.4 SST can be predicted from antecedent circulation and (Indian) oceanic departures during the late boreal winter, ten months before the peak phase of ENSO events. It is conjectured, that the regression models would yield a substantially inferior performance for prediction proper- that is, for years beyond the last year on which the regression model is based due to the large body of analyses already conducted on the selected precursors. However, this contention could not be verified from the short database used here. Nevertheless, the limited experiments conducted here do support the proposition that the Indian Ocean is now a crucial parameter in ENSO evolution. Thus, the inclusion of SEIO SSTs as a predictor may further improve the predictive skill of the current dynamical or statistical models which are currently used to forecast equatorial Pacific SST.

Acknowledgements: The BMRC upper Pacific Ocean 20°C isothern depth was kindly provided by Neville Smith. We thank Keith A. J. Rodgers for comments and careful re-reading of the manuscript.

References

Carton, J. A., G. Chepurin, X. Cao, and B. Giese (2000), A simple ocean data assimilation analysis of the global upper ocean 1950-95- Part I: Methodology, *J. Phys. Oceanogr.*, 30, 294-309

Clarke, A. J., and S. Van Gorder (2003), Improving El Niño prediction using a space-time integration of Indo-Pacific winds and equatorial Pacific upper ocean heat content, *Geophys. Res. Lett.*, 30, doi:10.1029/2002GL016673

Davison, A. C., and D. V. Hinkley (1997), Bootstrap methods and their application, Cambridge University Press, Cambridge, UK, pp 582

Goddard L., and N. E. Graham (1997), El Niño in the 1990s, *J. Geophys. Res.*, 102, 10423-10436

Jin, F. -F. (1997), An equatorial ocean recharge paradigm fo ENSO. Part I: conceptual model, *J. Atmos. Sci.*, 54, 811-829

Kalnay, et al. (1996), The NCEP-NCAR 40-year reanalysis project, *Bull. Amer. Meteo. Soc.*, 77, 437-471

Kug, J.-S., S.-I. An, F. F. Jin, and I.-S. Kang (2005), Preconditions for El Niño and La Niña onsets and their relation to the Indian Ocean, *Geophys. Res. Lett.*, 32, L05706.

Landsea, C. W., and J. A. Knaff (2000), How much skill was there in forecasting the very strong 1997-98 El Niño?, *Bull. Am. Meteorol. Soc.*, 81, 2107-2119

McPhaden, M. J., and Coauthors (1998), The tropical ocean global atmosphere observing system : A decade of progress, *J. Geophys. Res.*, 103, 14 169-14 240

Meinen, C. S., and M. J. McPhaden (2000), Observations of warm water volume changes in the Equatorial Pacific and their relationship to El Niño and La Niña, *J. Clim.*, 13, 3551-3559

Nitta, T., and S. Yamada (1989), Recent warming of tropical sea surface temperature and its relationship to the Northern Hemisphere, *J. Meteor. Soc. Japan*, 67, 375-383

Smith, N. R. (1995), An improved system for tropical ocean subsurface temperature analysis, *J. Atmos. Oceanic Technol.*, 12, 850-870.

Smith, T. M., and R. W. Reynolds (2004), Improved extended reconstruction of SST (1854-1997), *J. Clim.*, 17, 2466-2477.

Terray. P., S. Dominiak and P. Delecluse (2005), Role of the southern Indian Ocean in the transitions of the monsoon-ENSO system during recent decades, *Clim. Dyn.*, vol. 24, 169-195.

Terray. P., and S. Dominiak (2005), Indian Ocean Sea Surface Temperature and El Niño-Southern Oscillation: a new perspective, *J. Clim.*, vol. 18, 1351-1368.

Torrence C., and P. J. Webster (2000), Comments on “The Connection between the Boreal Spring Southern Oscillation Persistence Barrier and Biennial Variability”, *J. Clim.*, 13, 665-667

Trenberth, K., and T. Hoar (1996), The 1990-1995 El Niño-Southern Oscillation event: longest on record, *Geophys. Res. Lett.*, 23, 57-60

Van Loon, H., G. H. Meehl, and R. Milliff (2003), The Southern Oscillation in the early 1990s. *Geophys. Res. Lett*, 30, 311-314

Wang, B. (1995), Interdecadal changes in El Niño onset in the last four decades, *J. Clim.*, 8, 267-285

Wang, C., and R. H. Weisberg (2000), The 1997-98 El Niño Evolution Relative to Previous El Niño Events, *J. Clim.*, 13, 488-501

Figure captions

Figure 1: a) Correlation analysis of December-January Niño3.4 (5°S-5°N/190°E-240°E) SST time series with February-March 10-m zonal wind over the tropical Indian and Pacific Oceans ten months before for 1977-2001 period. The correlation coefficients significant at the 10% confidence level according to a phase-scrambling bootstrap test (Davison and Hinkley, 1997) are shaded. b) Same as a) but for the February-March mean depth of the 20°C isotherm. c) Same as a) but for the February-March SST fields.

Figure 2: October-December Niño3.4 SST, observed (solid line and closed squares) and forecast by the Z20 (dotted-dashed line and open circles), UWPAC (dashed line and grey-filled circles) and SEIO (dotted line and closed circles) univariate regression models without involving data for the forecast year. The horizontal dotted lines show the seasonal mean, mean minus one standard-deviation and mean plus one standard-deviation Niño3.4 SST values.

Figure 3: Same as figure 2) but for UWPAC-SEIO, Z20-SEIO and Z20-UWPAC bivariate regression models. UWPAC-SEIO model is dotted line and closed circles, Z20-SEIO model is dashed line and grey-filled circles and Z20-UWPAC is dotted-dashed line and open circles.

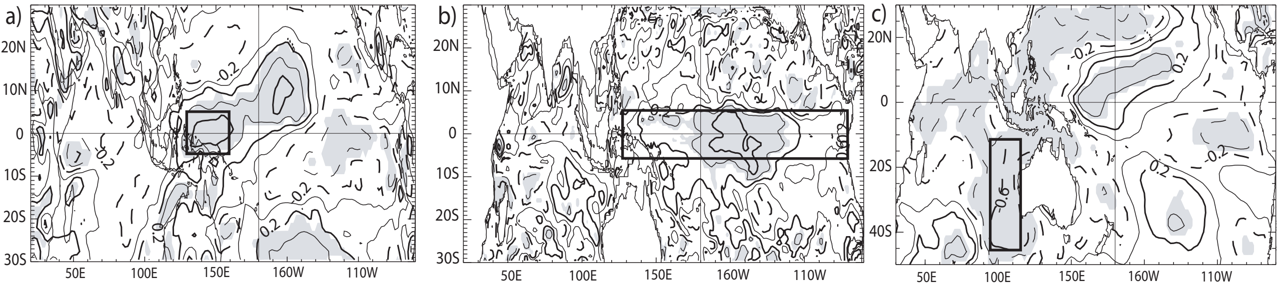
Table Captions

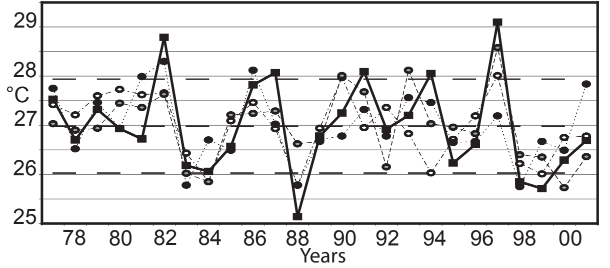
Table 1: Appraisal of Niño3.4 SST predictions from the February-March season for different lead times (10-12, 11-01, 12-02) around the peak phase of ENSO events by simple linear regression models using one predictor (UWPAC, Z20 or SEIO) as sole imput. The forecast skill of the models is assessed by cross-validated RMSE, COR and PP statistics calculated from the observed and forecast Niño3.4 SST by each model without involving data of the forecast year. Correlation coefficients significant at the 10%, 1%, 0.01% levels are indicated by one (*), two (**) and three stars (***), respectively.

Table 2: Same as Table 1, but for the two-predictors regression models using two elements as input from the set of UWPAC, Z20, SEIO precursors.

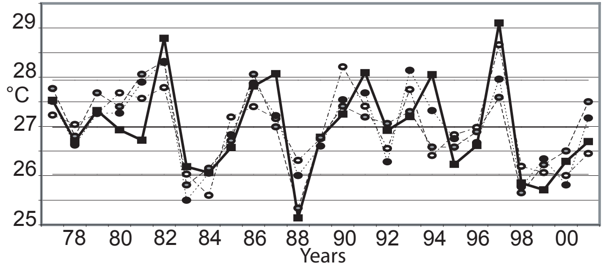
Table 3 : Same as Table 1, but for the three-predictors regression model using all the UWPAC, Z20, SEIO precursors.

Table 4 : Same as Table 1, but for the one and three-predictors regression models using the Z20 precursor computed from the BMRC reanalysis during the 1980-2001 period.


a)
b)
c)
20N
10N
0
10S
20S
30S
40S
50E
100E
150E
160W
110W
-0.2
0.2
-0.6



29
28
°C
27
26
25
78
80
82
84
86
88
90
92
94
96
98
00
Years


29
28
°C
27
26
25
78
80
82
84
86
88
90
92
94
96
98
00
Years

Forecast period	Regressors	RMSE	COR	PP
10-12	Z20	0.74	0.62 **	0.40
	UWPAC	0.67	0.70 ***	0.51
	SEIO	0.68	0.69 ***	0.50
11-01	Z20	0.78	0.63 **	0.42
	UWPAC	0.73	0.68 **	0.49
	SEIO	0.67	0.74 ***	0.57
12-02	Z20	0.77	0.62 **	0.40
	UWPAC	0.75	0.64 **	0.44
	SEIO	0.67	0.73 ***	0.55

Forecast period	Regressors	RMSE	COR	PP
10-12	Z20 - UWPAC	0.67	0.71 ***	0.52
	UWPAC - SEIO	0.58	0.79 ***	0.64
	Z20 - SEIO	0.64	0.73 ***	0.55
11-01	Z20 - UWPAC	0.71	0.71 ***	0.51
	UWPAC - SEIO	0.60	0.79 ***	0.65
	Z20 - SEIO	0.65	0.76 ***	0.59
12-02	Z20 - UWPAC	0.72	0.68 **	0.48
	UWPAC - SEIO	0.63	0.75 ***	0.60
	Z20 - SEIO	0.65	0.75 ***	0.57

Forecast period	Regressors	RMSE	COR	PP
10-12	SEIO - UWPAC - Z20	0.59	0.78 **	0.62
11-01		0.63	0.78 ***	0.62
12-02		0.65	0.76 ***	0.58

Forecast period	Regressors	RMSE	COR	PP
10-12	Z20_BMRC	0.77	0.62 **	0.42
	SEIO - UWPAC - Z20_BMRC	0.62	0.78 ***	0.62
11-01	Z20_BMRC	0.80	0.65 **	0.45
	SEIO - UWPAC - Z20_BMRC	0.64	0.80 ***	0.65
12-02	Z20_BMRC	0.77	0.67 **	0.47
	SEIO - UWPAC - Z20_BMRC	0.69	0.75 ***	0.58